DOCUMENT RESUME

ED 130 605 IR 004 133

AUTHOR	Shettel, Harris H.; Lindley, Richard H.
TITLE	An Experimental Comparison of Two Types of Self-Instructional Programs for a SAGE System Paired-Associate Task.
INSTITUTION	American Institutes for Research in the Behavioral Sciences, Pittsburgh, Pa.
SPONS AGENCY	Air Force Systems Command, L.G. Hanscom Field, Mass. Electronic Systems Div.
REPORT NO	AFCCDD-TN-61-20; AIR-C-11-61-TR-251
PUB DATE	Mar 61
CONTRACT	AF-19(604)-5951
NOTE	28p.; Archival document
EDRS PRICE	MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS	*Autoinstructional Programs; Comparative Analysis; Educational Research; *Paired Associate Learning; *Response Mode; Teaching Methods; *Time Factors (Learning)

ABSTRACT

A study was conducted to determine the optimal presentation methods for teaching the phonetic alphabet which is characteristic of much of the symbolic material which forms part of the SAGE L-system operator task. This discrete-item, paired-associate material was prepared in six different formats: (1) long continuous-discourse program, overt responses; (2) short continuous-discourse program, overt responses; (3) same as 2 but using covert responses; (4) flash cards with time limited, covert responses; (5) flash cards, time unlimited, covert responses; and (6) flash cards, time unlimited, with drop-out of mastered items. Each program was given to ten undergraduate subjects. Procedures and results were discussed in detail, and tabular summaries are provided. (Author/EMH)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

UNCLASSIFIED

AFCCDD-TN-61-20

AIR-C11-61-TR-251

ED130605

AN EXPERIMENTAL COMPARISON OF TWO TYPES OF SELF-INSTRUCTIONAL PROGRAMS FOR A SAGE SYSTEM PAIRED-ASSOCIATE TASK

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Harris H. Shettel
Richard H. Lindley

AMERICAN INSTITUTE FOR RESEARCH
Pittsburgh 32, Pennsylvania

MARCH 1961

Contract No. AF 19(604)-5951

Prepared for
OPERATIONAL APPLICATIONS OFFICE
DEPUTY FOR TECHNOLOGY
ELECTRONIC SYSTEMS DIVISION
AIR FORCE SYSTEMS COMMAND
BEDFORD, MASSACHUSETTS

IR 004 133

ACKNOWLEDGMENTS

The authors wish to extend their appreciation to Arthur A. Lumsdaine and Robert Glaser for the helpful suggestions they made during the course of this study. Thanks are also due to Richard B. Ross for his assistance in running subjects.

FOREWORD

This is one of a series of seven reports prepared by the American Institute for Research under Contract AF 19(604)-5951, Development of optimal methods of organizing and presenting SAGE operator job information in self-training equipment. The complete series is listed below, with authorship, title, and report number indicated for each document.

1. Shettel, H. H., Angell, D., Fazio, Janet L., Lindley, R. H. & Lumsdaine, A. A. Training requirements for the development of prototype self-instructional programs covering basic job knowledge for SAGE Track Monitor and Intercept Director operator positions. Pittsburgh: American Institute for Research, 1961. AIR-C11-61-SR-246 (ESD-TN-61-52)

2. Shettel, H. H. & Lumsdaine, A. A. Principles of programming as applied to the development of two self-instructional programs for SAGE operators. Pittsburgh: American Institute for Research, 1961. AIR-C11-61-SR-247 (AFCCDD-TN-61-27)

3. Shettel, H. H., Angell, D. & Lumsdaine, A. A. Handbook for the preparation of self-instructional materials for SAGE operators. Pittsburgh: American Institute for Research, 1961. AIR-C11-61-SR-248 (AFCCDD-TN-61-30)

4. Shettel, H. H. & Lumsdaine, A. A. Characteristics of self-instructional devices to meet training requirements for the SAGE Track Monitor, Intercept Director, and related system-operator positions. Pittsburgh: American Institute for Research, 1961. AIR-C11-61-SR-249 (ESD-TN-61-47)

5. Shettel, H. H. Evaluation of two prototype self-instructional programs covering basic job knowledge for SAGE Track Monitor and Intercept Director operator positions. Pittsburgh: American Institute for Research, 1961. AIR-C11-61-SR-250 (ESD-TN-61-51)

6. Shettel, H. H. & Lindley, R. H. An experimental comparison of two types of self-instructional programs for a SAGE system paired-associate task. Pittsburgh: American Institute for Research, 1961. AIR-C11-61-TR-251 (AFCCDD-TN-61-20)

7. Shettel, H. H., Angell, D. & Lumsdaine, A. A. FINAL REPORT (Short title) Self-instructional programs for SAGE system operators. Pittsburgh: American Institute for Research, 1961. AIR-C11-61-FR-252 (ESD-TR-61-17)

The project staff wishes to acknowledge the cooperation of the military agencies and personnel contacted, and the close support provided by Dr. Walter Organist and Dr. Sylvia Mayer, of the monitoring agency. Captain Raigh Mason and Mr. Ira Goldstein, also of the monitoring agency, were particularly helpful in assisting in the conduct of the field evaluation of the completed programs.

Members of the research staff of the American Institute for Research who made major contributions to various phases of the work done under this contract include: Arthur A. Lumsdaine (Principal Investigator), Harris H. Shettel, David Angell, Janet Fazio, and Richard Lindley.

5

ABSTRACT

A problem in implementing self-instructional methodology for the training of SAGE and other L-system operators is the wide variety of behaviors and skills required by these operators and the wide variety of training techniques available to the self-instructional programmer for teaching these skills and behaviors. The study was done to determine optimal presentation methods for teaching the phonetic alphabet which is characteristic of much of the symbolic material which forms a part of L-system operator tasks. This discrete-item, paired-associate material was prepared in six different formats: (1) long continuous-discourse program, overt responses; (2) short continuous-discourse program, overt responses; (3) same program as (2) but using covert responses; (4) flash cards, time limited, covert responses; (5) flash cards, time unlimited, covert responses; and (6) flash cards, time limited, with drop-out of mastered items. Each program was given to ten college undergraduate subjects.

The least number of errors on two immediate post-tests were made by the subjects who were given the long program and any of the flash card programs. The most errors were made by subjects receiving either short program.

The three flash card groups led to high levels of performance in short periods of time (14 minutes) as compared with the long, overt response program (40 minutes). Thus, when preparation time, training time, and amount learned are all considered, conventional self-instructional programming methods appear to have no advantage over simpler presentation techniques for teaching discrete-item, paired-associate SAGE subject matter.

TABLE OF CONTENTS

	Page
FOREWORD	iii
ABSTRACT	v
INTRODUCTION	1
PROBLEM	3
METHOD	3
Experimental design	3
Materials	5
Procedure	8
Subjects	8
RESULTS	8
Training times	9
Proficiency tests	10
DISCUSSION	13
IMPLICATIONS	16
REFERENCES	21

List of Tables

Table 1. Experimental Design	6
Table 2. Sample Frames from Constructed Response Programs	7
Table 3. Median Training Times and Ranges	9
Table 4. Number of Ss in Each Group who were Below and Above the Overall Median Training Time	10
Table 5. Error Data for All Groups	11
Table 6. Efficiency Scores and Ranks for All Groups	12
Table 7. Values Assigned to Each Training Group in Terms of Three Parameters	18

INTRODUCTION

Those who are engaged in the preparation of self-instructional programs face a practical problem in deciding on the best methods or techniques to use for teaching various kinds of subject matter. This problem is particularly important when the task being programmed is complex, including a variety of skill and knowledge areas. Research evidence is accumulating which supports the notion that a single self-instructional technique will not be found to be optimally effective for all subject matter areas but rather that specific approaches will be found to work best for certain kinds of subject matter (Coulson & Silberman, 1959; Fry, 1960; Goldbeck, 1960).

In the course of developing self-instructional training programs for the SAGE Track Monitor and Intercept Director operator positions it became apparent that some of the symbolic subject matter was awkward to program in the conventional sense, i.e., continuous-discourse, small step, overt write-in frames. This symbolic material was characterized as being paired-associate in nature, with very little internal or external referability between the various pairs of items. Thus, the phonetic alphabet, as an example, consists of the 26 letters of the alphabet which are paired with 26 code words. There is no necessary connection between the "A -- alfa" and "B -- bravo" pairs; they are discrete items. Neither is there any particular significance to the fact that a track is called "Mike 407" rather than "Oscar 407." The only requirement in learning the phonetic alphabet is that the operator be able to translate the visual image "M407" on his visual display into the auditory response "Mike 407," or conversely, that he can find "M407" on his display in response to the auditory stimulus "Mike 407." The kind of control exerted over the learning situation by conventional self-instructional methods seemed inappropriate to the simplicity of this kind of subject matter.

In contradistinction to this discrete-item material is another kind of paired-associate material, characterized by track symbology. Track symbology is not discrete but rather related internally to other symbols and externally to specific operator actions that would be taken in response to various combinations of these track symbols. This type of symbolic material is closely

related to concept formation material, and for this reason was quite amenable to the kind of control (or careful shaping of responses) characterized by the programming techniques used in the main body of the self-instructional training programs.

In addition to the question of teaching effectiveness, there is the question of programming effort. Perhaps complex programming techniques <u>can</u> teach discrete-item, paired-associate material. But if simpler methods of presenting such material work as well, their use would be recommended on purely practical grounds, particularly in view of the fact that a large amount of such material is part of basic job knowledge for SAGE, and other L-system operators.

Thus, the study reported here grew out of both practical and theoretical considerations. The more detailed the control exerted over the learning situation, the more effort required by the programmer. The question this study was aimed at answering is: What combination of techniques would provide for the most efficient learning of a discrete-item, paired-associate SAGE task with the least expenditure of programming time and effort?

PROBLEM

Self-instructional programs have been compared with the tutorial method of teaching (e.g., Lumsdaine, 1959). Both the self-instructional and tutorial methods attempt to carefully shape the responses of S and to reinforce responses by providing immediate knowledge of results. A basic notion is that the instructor, be it tutor or program, facilitates learning by controlling the behavior of S more closely than is usually possible with other methods of teaching.

The present study compared the amount of learning produced by various types of programs for teaching discrete-item, paired-associate material. The material taught was the phonetic alphabet in which each letter of the alphabet has an assigned code word (e.g., A = ALFA, B = BRAVO). Three versions of continuous-discourse programs elicited responses throughout the training period, provided knowledge of results soon after the responses were made, and provided practice in a task very similar to one used as a proficiency test at the end of training. Three other programs consisted basically of sets of flash cards in which a letter of the alphabet appeared on one side of the card, the appropriate code word appeared on the other side, and S was instructed to learn the paired associates by the anticipation method. Each version of the program represented a different level of control over the learning behavior of the subject. The purpose of the present study was to determine which type of program led to more effective and more efficient learning.

METHOD

Experimental design

There were three groups trained with the continuous-discourse programs and three groups trained with the flash cards. In the description of groups below, the expression "training time limited" means that E and not S determined the amount of time that could be spent with the training materials.

Group 1, long constructed response program, overt responding, training time not limited. This group was trained with a long constructed response program consisting of 90 frames. The Ss responded overtly by writing the appropriate responses in the spaces provided.

Group 2, short constructed response program, overt responding, training time not limited. This group was trained with a short constructed response program consisting of 30 frames selected from the long program. The Ss responded overtly by writing the appropriate responses in the spaces provided.

Group 3, short program, covert responding, training time not limited. This group was trained with the same short 30 frame constructed response program as Group 2 but was instructed to simply think of the correct response every time a space for a response appeared in the program (i.e., covert responding); the Ss did not write during training. A comparison of the performance of Groups 2 and 3 would show whether overt or covert responding produces differences in the amount learned for this task.

Group 4, flash cards, covert responding, training time limited. This group was trained with the flash cards using the anticipation method and was instructed to look at the letter of the alphabet which appeared on one side of the card and to try to remember the phonetic equivalent (after the first trial) before looking at it on the other side. The time allotted the Ss was determined by pairing each S in Group 4 with an S from Group 2; the amount of time taken by the S in Group 2 to go through the constructed response program with overt responding determined the amount of time allotted his matched S in Group 4.

Group 5, flash cards, covert responding, training time not limited. This group was trained with the flash cards and received the same instructions about the anticipation method as were received by Group 4. Group 5 differed from Group 4 in the amount of time the Ss practiced with the flash cards; Ss in Group 5 were instructed to practice until they reached a criterion of two successive errorless trials and then to notify E. As soon as S notified E, the training was stopped.

Group 6, flash cards, covert responding, "drop-out" feature, training time limited. This group was trained with the flash cards and received the same instructions about the anticipation method as Groups 4 and 5. In addition, however, the Ss were instructed to sort the flash cards into two piles during learning -- one containing the phonetic equivalents that they had anticipated correctly and one containing the equivalents that they had not anticipated correctly. The Ss were then to practice with the cards in the latter pile; once again placing cards for which they made correct responses in the proper pile. In this way cards containing phonetic equivalents that S had anticipated correctly were temporarily dropped out of the program and S practiced with only those equivalents that he did not know. When every code word had been correctly anticipated once, S then practiced with a set of practice cards containing short words for which S was to think of the phonetic equivalent for each letter. The correct answers were on the reverse side of the cards. When S completed the set of practice cards, he returned to the flash cards and repeated the sequence of anticipating the code words, sorting the cards into piles, practicing with the equivalents he had not anticipated correctly, and finally, practicing spelling words phonetically. The training times allotted the Ss in this group were the same as for Groups 2 and 4; each S was matched with a corresponding S from Group 2.

A summary of the experimental design is presented in Table 1. The first column of this table refers to whether or not the training program contained practice in a task that comprised one of the two tests used after training.

Materials

The flash card programs consisted of 26 cards. A letter of the alphabet was typed on each card and its phonetic equivalent was typed on the reverse side. For Groups 4 and 5 the cards were arranged in alphabetical order in a bound booklet. The cards for Group 6 were unbound to permit S to sort the cards into separate piles.

The practice cards (as distinct from the training cards) for Group 6 consisted of 12 cards. A short English word was typed on one side of each card and the phonetic equivalents of the letters were typed on the other side (thus, "RED -- romeo, echo, delta"). The 26 letters of the alphabet appeared at least once in

Table 1

Experimental Design

Groups	Training Variables*		
	Phonetic Spelling Practice Present (P) vs Absent (A)	Limited (L) vs Unlimited (U) Training Time	Overt (O) vs Covert (C) Responding
1 Long continuous-discourse program	P	U	O
2 Short continuous-discourse program	P	U	O
3 Short continuous-discourse program	P	U	C
4 Flash cards	A	L	C
5 Flash cards	A	U	C
6 Flash cards**	P	L	C

*The letters in the columns indicate which variable was present, e.g., the "P" in the first column for Group 1 shows that phonetic spelling practice was present during training for this group.

**Group 6 also was trained with the "drop-out" feature described in the text.

the words used for the practice cards; none of these same words was used in one of the two proficiency tests given at the end of training which required S to similarly spell words phonetically.

The long constructed response program consisted of 90 frames; the two short programs consisted of 30 frames selected from the long program. Three sample frames are presented in Table 2. In general the sequence was a frame requiring copying responses, two or three frames requiring fill-in responses, another frame of copying responses, and then frames containing words to be phonetically spelled (as described above). As proficiency with any given series of letters increased,

Table 2

Sample Frames from Constructed Response Programs

1. Copying Response Frame (S responds in boxes: [ ]).

   The first five letters of the alphabet are A = Alfa, B = Bravo, C = Charlie, D = Delta, E = Echo. Therefore, E is [ ], C is [ ], A is [ ], D is [ ], and B is [ ].

2. Fill-in Response Frame (S fills-in blank spaces).

   A Al_a

   C Cha_ie

   B _ravo

   D Del_a

   E Ech_

3. Phonetic Spelling Frame (S responds in boxes: [ ]).

   In the phonetic alphabet C A B is Charlie, Alfa, Bravo, and B E D is Bravo, Echo, Delta. B A D is [ ] [ ] [ ] and C A D is [ ] [ ] [ ].

the program reduced the cues for the correct response that were available to S. For example, the first time a series of letters was introduced, the first frame required copying responses with the correct answers available on the frame; the second time these responses were required the correct responses were not in view of S as he responded.

The two short and one long continuous-discourse programs differed principally in the number of frames requiring phonetic spelling of words; both types of programs had 10 frames of copying responses, 10 and 11 frames respectively of fill-in responses, and 10 and 67 frames respectively of responses requiring the phonetic spelling of words. The long program also had two review frames (one in the middle of the program and one at the end) which required fill-in responses. The correct answers appeared on the back of each frame; a blank page was placed between all

frames to prevent S from seeing answers from the previous frame.

Two tests were used at the end of training, a straight recall test requiring S to write the entire phonetic alphabet and a phonetic spelling test. The latter consisted of 26 words typed on separate cards in a bound booklet. S looked at the word and then spelled it aloud phonetically. Each letter of the alphabet appeared at least three times in the test words used.

Procedure

The appropriate instructions were first read to S and then he was given one of the six programs. The Ss in Groups 1, 2, 3, and 5 were allowed as much time as they needed to go through their programs; these Ss were assigned to groups using a randomized block method. After each S in Group 2 had been tested, an S was assigned to Group 4 and another S to Group 6; the latter two Ss were allotted the same amount of time during training as was taken by the matched S in Group 2.

At the end of training the Ss were given both tests, first the recall test and then the phonetic spelling test. S was told that he would be scored for both accuracy and speed. In both tests a failure to respond or an incorrect response constituted an error. In the phonetic spelling test repeating an error on the same letter did not add to S's error score; an error was counted in this test if S made one or more errors on a given letter.

Subjects

The Ss were 60 undergraduate men and women from Carnegie Institute of Technology. None of the Ss had had previous experience with the phonetic alphabet. Ten Ss were assigned to each group.

RESULTS

The analysis of variance technique was not used since the variances of the error scores were highly heterogeneous. Instead, a nonparametric X^2 (chi-square) test was used to test the statistical significance of both the error and time scores.

Training times

The median training times and ranges to complete the programs are presented in Table 3. The X^2 median test was used to test the significance of the differences in training times among the groups; the first step in the test was to compute the overall median training time for all Ss irrespective of the group to which the Ss had been assigned. The number of Ss in each group who were below and above the overall median training time are presented in Table 4; the second step in the X^2 median test was to compute X^2 for this 2 x 4 contingency table. The X^2 computed from the data in Table 4 was 21.60, $P < .01$, df = 3. Thus the differences in training times for the groups were highly significant; inspection of Tables 3 and 4 reveals that the training times for both the long and short continuous-discourse programs with overt responding (Groups 1 and 2) were considerably longer than the training times for both the short continuous-discourse program with covert responding (Group 3) and the Group 5 flash card program.

Table 3

Median Training Times and Ranges

Groups	Median	Ranges
1	40 min. 02 sec.	28 min. 07 sec. -- 50 min. 19 sec.
2*	16 min. 18 sec.	13 min. 39 sec. -- 21 min. 53 sec.
3	10 min. 41 sec.	7 min. 50 sec. -- 15 min. 29 sec.
5	11 min. 50 sec.	3 min. 39 sec. -- 21 min. 21 sec.

*Groups 4 and 6 received the same training times as Group 2.

Table 4

Number of Ss in Each Group who were Below and Above the Overall Median Training Time

Group	Below Median	Above Median
1	0	10
2*	3	7
3	9	1
5	8	2

*Groups 4 and 6 same as Group 2.

These results are hardly surprising; they show that the long overt continuous-discourse program takes longer to complete than the short overt continuous-discourse program and that the latter takes longer to complete than the short covert continuous-discourse program. However, these results are important when test error scores are considered as well as the length of the training period.

Proficiency tests

Table 5 presents the median and mean numbers of errors and the ranges for both the recall and phonetic spelling tests. Both tests showed that the Ss who were taught with the flash cards performed better than the Ss who were taught with the two short continuous-discourse programs. Considering all three of the continuous-discourse programs together, the recall test error scores revealed there were 13 Ss below the median and 17 Ss above it; in contrast there were 25 Ss in the three flash card groups below the median error score and only 5 Ss above it. These differences are significant ($X^2 = 10.34$, $P < .01$, $df = 1$). (Owing to the fact that a large number of Ss received the median error score, the data could not be classified so that exactly 50 percent of the Ss were below the median and exactly 50 percent above it. Therefore, the Ss were classified so that the number of Ss above and below the median was as close to 50 percent as possible. This classification presented no problems in the statistical analysis as the expected frequencies for X^2 were adjusted to conform to the actual classification. This procedure is described in Guilford, 1956, Chapter 11.)

Table 5

Error Data for All Groups

Groups	Recall Test Errors			Phonetic Spelling Test Errors		
	Means	Medians	Ranges	Means	Medians	Ranges
1	0.30	0.00	0-2	1.00	1.00	0-4
2	3.80	2.50	0-12	6.00	5.00	0-17
3	2.50	1.50	0-10	4.30	2.00	0-17
4	0.10	0.00	0-1	0.60	0.00	0-4
5	0.20	0.00	0-1	3.30	1.50	0-10
6	0.30	0.00	0-2	1.10	0.00	0-7

The phonetic spelling test proved less sensitive than the recall test although the results still reached statistical significance. In the phonetic spelling test there were 9 Ss in the three continuous-discourse program groups below the median and 21 above it; there were 17 Ss in the three flash card groups below the median and 13 above it ($X^2 = 4.34$, $P < .05$, $df = 1$). Thus, as groups, the flash card programs were more effective than the continuous-discourse programs.

When the performance of the individual groups is considered, only the error data of the recall test reached an acceptable level of statistical significance; therefore, the error data of the phonetic spelling test will not be presented here. The error data of the phonetic spelling test did, however, show the same trends as the error data of the recall test.

The number of Ss below the median for the recall test error data for each

group was as follows: Group 1, 8; Group 2, 3; Group 3, 2; Group 4, 9; Group 5, 8; Group 6, 8 ($X^2 = 19.51$, $P < .01$, df = 5). This analysis clearly shows that the two short continuous-discourse programs (either covert or overt responding), i.e., Groups 2 and 3, led to the production of more errors than any of the other programs.

Time to complete the writing of the phonetic words in the recall test was also measured. The number of Ss below the median testing time for the recall test for each group was as follows: Group 1, 8; Group 2, 2; Group 3, 2; Group 4, 8; Group 5, 6; Group 6, 4 ($X^2 = 15.20$, $P < .01$, df = 5). Thus the longest testing times for the recall test were for the short continuous-discourse programs with either covert or overt responding. Not only did the short continuous-discourse programs lead to more errors, but they also led to longer testing times; the groups that made the fewest errors also took the least time to complete the test.

Scores were also computed for all groups based on the number of items correct on the post-tests (Table 5) divided by learning time (Table 3). This score provides an index of the efficiency of the various instructional techniques as they were used in this study, i.e., how much they taught per unit of learning time. The data from this analysis are presented in Table 6.

Table 6

Efficiency Scores and Ranks for All Groups

Groups	Efficiency Score	Efficiency Rank
1	.63	6
2	1.31	5
3	2.12	1
4	1.57	3
5	2.05	2
6	1.55	4

Group 3 was the most efficient, with the three flash card groups following in rank. The two constructed response programs (Groups 1 and 2) were the least efficient.

DISCUSSION

The error data clearly show that the flash card programs produced fewer errors than did the short continuous-discourse programs with either overt or covert responding. It would be of considerable importance to discover the particular variables that led to the generally superior performance of the flash card groups. One variable may have been the temporal interval between making a response and receiving knowledge of results. In the flash card programs each response was followed by knowledge of results as soon as S turned over the card. In the continuous-discourse programs knowledge of results was often delayed until S had made several responses on the same frame.

A second variable which may account for the difference is the size of the unit of material that was taught. The Ss in the flash card groups could practice with as many or as few paired-associates as they wished; the Ss in the constructed response programs were restricted to practicing with only the material presented on a given frame. The continuous-discourse programs used in the present study may not have selected the optimal size of the unit of material to be learned; this variable clearly needs more extensive investigation (cf. Barlow, 1960). Of course, the individual S may ultimately turn out to be the best determiner of the size of the instructional unit for this type of material.

Requiring fill-in responses may also have prejudiced the continuous-discourse programs. While this technique is apparently appropriate for some subject matter as, for example, spelling as has been demonstrated by Skinner (1958), it may have been inappropriate for the learning of complete terms where spelling was not part of the criterion task.

All of these possible negative aspects (of the short programs) may have been attenuated by the amount of sheer practice given S in the long overt program. Its disadvantage was not in the level of learning attained, but in the time spent in reaching that level.

It was anticipated that the flash card program with the "drop-out" feature (Group 6) would produce superior learning as compared to the other flash card groups due to the fact that the former Ss should have received the most practice with the items with which they had the most difficulty. The results gave no support for this hypothesis; however, it seems likely that the Ss in the flash card groups without drop-out may have utilized this technique anyway by simply rehearsing most with the items with which they had difficulty. Less motivated and/or less sophisticated subjects may not produce similar results.

The present study showed that overt responding with the continuous-discourse program leads to longer training times than covert responding without producing a compensating reduction in errors. Kaess and Zeaman (1960) also found no difference in performance when covert and overt responding were compared. These results suggest that programmers should continue to investigate the use of covert responding in programmed teaching.

The present study also illustrates the possibility of "over-programming" discrete-item, paired-associate material. In preparing a program which presented the material in very small steps, with attention paid to the systematic and gradual reduction of cue strength, it was easy (or so it was found) to write what appeared to be an excessively long program. Unfortunately, it was not possible to obtain long-term recall data (e.g., an interval of one month or so) from the Ss who participated in this study; perhaps the continuous-discourse programs would have demonstrated an advantage over the simpler techniques in the long-term recall of discrete-item, paired-associate material.

The efficiency data presented in Table 6 must be interpreted with caution. It does indicate how efficiently the particular techniques used in the study reached their respective levels of achievement. It does not, however, allow

one to extrapolate from the data so that general statements of efficiency can be made concerning the value of these techniques under other conditions and/or for other achievement levels. For example, the short covert program could be recommended as the most efficient only if the errors made by this group are acceptable. It is not possible to assign the same efficiency to a similar program designed to raise the level of learning unless an assumption of linearity can be made for the learning curve. This assumption is probably unwarranted for paired-associate material where the learning curves are usually negatively accelerated as they approach asymptote. To the extent that this is true, efficiency scores for low achievement groups (such as Group 3) would be spuriously high. Thus, it may (and probably does) require a disproportionately longer time to teach the final few paired-associate items in the phonetic alphabet than it does to teach earlier items.

To state the problem in more general terms, if it requires 15 minutes to teach 15 items, using a specific technique, it does not necessarily follow that it requires 26 minutes to teach 26 items using the same technique, an assumption which is implicit in using number right divided by training time as an overall efficiency score. Thus, unless learning is carried out to a set, or agreed upon criterion level (preferably a maximum level), or unless the learning curves can be demonstrated to be linear, the imputing of a certain degree of efficiency to a technique (or media), rather than to an instance of its use, is misleading.[1]

In conclusion, Group 3 did not reach acceptable levels of achievement and it was not felt that the apparent efficiency of the short covert program would be maintained if it were modified so as to produce achievement levels comparable with those obtained with the easier to prepare flash card materials.

---

[1] If the learning curve were nonlinear but the specific shape of the learning curve were known, it would be possible to extrapolate efficiency scores appropriately adjusted for the effect of the shape of the curve. Only a linear curve, however, would enable a simple one to one relationship to obtain.

Practical decisions as to which technique will provide acceptable levels of learning within acceptable levels of time and with acceptable expenditures of effort can only be made by weighing values assigned to each parameter, and cannot be made only on the basis of error scores, time scores, or even necessarily, as noted here, combinations of the two (Edwards, 1956).

IMPLICATIONS

The study reported here was designed to provide practical guidance to SAGE programmers faced with the problem of optimizing presentation techniques as they are applied to discrete-item, paired-associate subject matter (i.e., phonetic alphabet, brevity code, and certain digital symbology), in terms of (1) acceptable levels of learning, (2) minimum training time, and (3) minimum programming effort.

Considered in this practical context, the results could be interpreted as follows:

1. The long continuous-discourse program did a very effective job of teaching the material. It required considerable time on the part of the subjects to use and on the part of the programmers to prepare.

2. The short program with overt responding was the least effective teaching program, although it was less demanding of time for the subjects and the programmers.

3. The covert version was slightly better than the overt program as concerns its teaching effectiveness, although it was still at a low level. Its main advantage was in training time, which was shorter than for any other group.

4. The flash card technique under timed conditions was a very effective method of teaching the phonetic alphabet. It

was much easier to prepare this material than any of the previous programs. Training time was the same as for Group 2.

5. The untimed flash card materials were also very easy to prepare but not as effective in teaching ability as (4) above. Training time was very short.

6. The introduction of the drop-out feature for mastered items resulted in a flash card segment almost as effective as (4) in teaching ability, but it required a bit more effort to prepare the material. Training time was the same as for (2) and (4).

The qualitative statements above reflect the quantitative data. They do not say which technique is best, but simply verbalize the data in terms of the three parameters considered relevant in this instance: (1) effectiveness, (2) training time, and (3) programming time. The final decision as to which technique to use must rest on the values attached to these three parameters. Each situation would present a unique set of conditions. In the present instance the following statements reflect the values attached to each parameter:

1. Instructional effectiveness.

   It was considered very important that the task be learned to a high level, i.e., that errors on the tests be reduced to a minimum. Eighty percent or even 90% learning would be considered only barely acceptable.

2. Training time.

   Not as important a consideration as (1). Of some importance, however, in training materials that were otherwise quite long. Certainly length had no intrinsic virtue and would be tolerated only if it produced consequent high levels of learning not to be achieved with shorter programs.

3. Programming effort.

In the present instance, programming time and effort was quite important. Time to complete the programs was limited and the amount of paired-associate material was considerable. While not as important as teaching effectiveness, some of the trainees' time would be sacrificed for savings in programming time in this instance. (Note that for nonprototype programs, and where large numbers of subjects would be involved over extended time periods, this parameter may rightly be considered less important than training time.)

Table 7 shows the various training groups ranked according to how well they met each parameter and with an arbitrary value loading given to each parameter, based on the above discussion of the relative importance of each parameter in the present study. Thus Instructional Effectiveness is given a value of "3" (most important), Training Time "1" (least important), and Programming Effort "2" (fairly important).

Table 7

Values Assigned to Each Training Group in Terms of Three Parameters

	Instructional Effectiveness		Training Time		Programming Effort		Average	
Group	Rank	x Value=3	Rank	x Value=1	Rank	x Value=2	Rank	Rank x Value
1	5	15	1	1	1	2	2.3	6.0
2	1	3	3	3	2.5	5	2.2	3.7
3	2	6	6	6	2.5	5	3.5	5.7
4	5	15	3	3	5.5	11	4.5	9.7
5	3	9	5	5	5.5	11	4.5	8.3
6	5	15	3	3	4	8	4.0	8.7

The first group (long overt program) was a very effective instructional segment, and is thus given a top rank of "five" in that category. In terms of training time and programming effort this segment was assigned the lowest rank of "one." However, since instructional effectiveness was considered more "valuable" than the other parameters, Group 1 received a higher average rank x value score than either Groups 2 or 3, both of which had higher ranks in the "time" and "effort" categories.

For the (arbitrary) values assigned to the six training methods used in this study, the flash card groups clearly had the most desirable combination of features as indicated in the "Average" column. Other values, in other situations, of course, could change the order of preference. (Note that the same programs in other situations would not change their ranks, just the values assigned to the various parameters. New programs or new versions of the present programs may also change ranks.)

There was no clear advantage to the drop-out feature used in Group 6. Groups 4 and 5 differed (in technique) only in the length of training time. The best combination of features seemed to be scheduled training time with simple flash card presentation.

This technique was thus adopted and used in both the Intercept Director (IND) and Track Monitor (TM) SAGE self-instructional programs. Field trials of the completed IND program made it possible to check the results of the flash card materials for the phonetic alphabet, since one of the sections of the IND proficiency test required the subjects to write the entire alphabet. Four naive subjects (those who had no previous knowledge of SAGE or the phonetic alphabet) scored respectively 22, 26, 25, and 26 items correct on this test, for an average error rate of 1.3. This error rate, while not as good as that obtained in the study for the flash card groups, recall test, was better than that obtained for the short overt or covert continuous-discourse programs. As an additional consideration, the IND proficiency test was given an average of <u>four days after</u> the phonetic alphabet flash card segment had been used in the program; the four days included interpolated activity consisting in part of drill with

similar paired-associate symbolic material.

It should be pointed out that the field trials noted above were conducted under fully monitored conditions. There is no assurance that the flash card technique would be equally effective when used in unmonitored situations. While it may be anticipated that highly motivated trainees would follow directions and use the materials properly, it cannot safely be assumed that less highly motivated subjects would do so.

In summary, not only did this study point to the flash card groups as representing the most effective combination of factors, but the application of this technique in the field under monitored conditions supported the research findings. Where similar values and conditions apply, similar flash card presentation techniques can be recommended for the training of discrete-item, paired-associate material of the type used in the present study.

REFERENCES

Barlow, J. Project tutor. Psychol. Rep., 1960, 6, 15-20.

Coulson, J. E. & Silberman, H. F. Results of initial experiment in automated teaching. In A. A. Lumsdaine & R. Glaser (Eds.) Teaching machines and programmed learning: A source book. Washington, D. C.: National Education Association, 1960. (Pp. 452-68)

Edwards, W. D. The use of statistical decision functions in making practical decisions. In G. Finch & F. Cameron (Eds.) Symposium on Air Force human engineering, personnel, and training research. Washington, D. C.: National Academy of Sciences -- National Research Council, 1956.

Fry, E. B. A study of teaching machine response modes. In A. A. Lumsdaine & R. Glaser (Eds.) Teaching machines and programmed learning: A source book. Washington, D. C.: National Education Association, 1960. (Pp. 469-74)

Goldbeck, R. A. The effect of response mode and learning material difficulty on automated instruction. Santa Barbara, Calif.: American Institute for Research, 1960.

Guilford, J. P. Fundamental statistics in psychology and education. New York: McGraw-Hill, 1956.

Kaess, W. & Zeaman, D. Positive and negative knowledge of results on a Pressey-type punchboard. J. exp. Psychol., 1960, 60, 12-17.

Lumsdaine, A. A. Partial and more complete automation of teaching in group and individual learning situations. In E. Galanter (Ed.) Automatic teaching: The state of the art. New York: Wiley & Sons, 1959.

Skinner, B. F. Teaching machines. Science, 1958, 128, 969-77.